

# Aquisição e Armazenamento de Metainformação no Contexto de um Arquivo

**Miguel Ferreira**  
mferreira@dsi.uminho.pt

*Co-Autoria*  
**José Carlos Ramalho**  
jcr@di.uminho.pt

**XATA '04**  
13 de Fevereiro de 2004



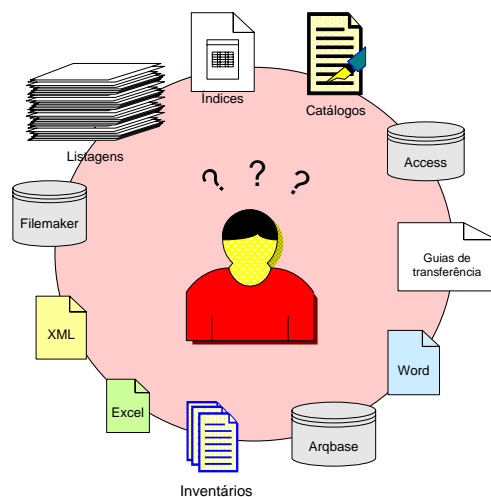
## Conteúdo

---

- Apresentação do **projecto**
- **Metainformação** de Arquivo
- Utilização de **XML**
- **Armazenamento** de metainformação
- **Avaliação** do modelo
- **Conclusões** e trabalho futuro



## Arquivo Distrital do Porto [ADP]



3



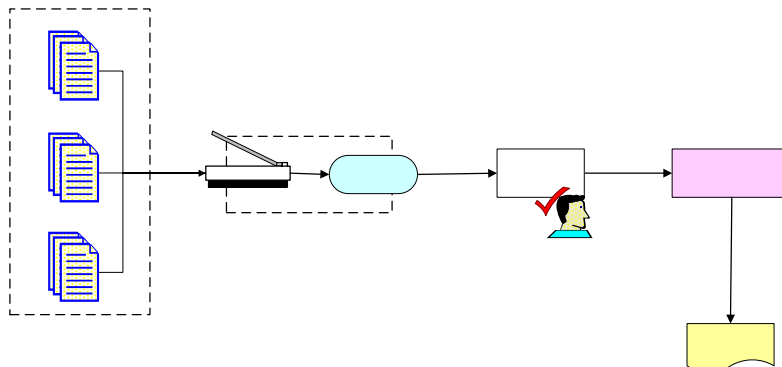
## Projecto DigitArq [objectivos]

- Centralização da informação
- Utilização de **normas** internacionais
  - *International Standard Archival Description* - ISAD(g)
  - *Encoded Archival Description* - EAD/XML
- Eliminação do **papel**
- Disponibilização do catálogo na **Web**

4



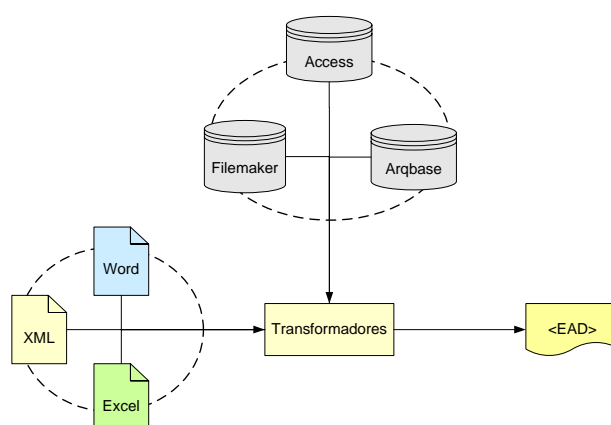
## Materiais em papel [digitalização]



5



## Bases de dados diversas [migração]



6

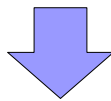


OCR

ação

## Resultado da migração

- Centenas de **ficheiros** EAD/XML
- **Gerir** toda a metainformação
- **Pequisar** e consultar via Web



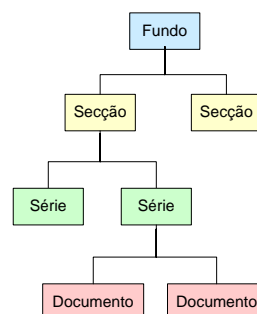
Como **armazenar**?

7



## Metainfomação Arquivística

- Estrutura **hierárquica**
- Descrição do mais **geral** para o mais **específico**
- **Níveis** de descrição
- Referência, título, datas extremas, condições de acesso, âmbito e conteúdo, etc.



8



## O XML neste contexto

### ■ Vantagens

- Modelo **hierárquico** intrínseco
- **Portabilidade** – fácil de trocar

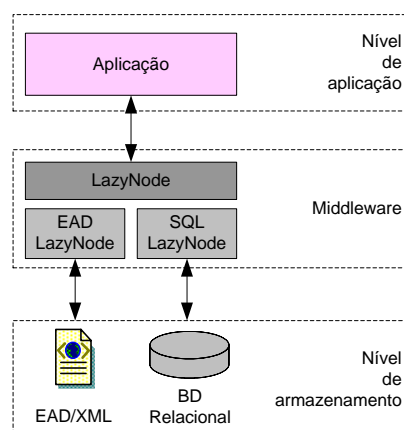
### ■ Desvantagens

- **Baixo nível**
- Difícil de **armazenar**

9



## Armazenamento de meta-informação



- Download()
- Upload()
- Children()
- CreateNode()
- AppendChild(child)
- RemoveChild(child)
- HasChildren()
- Clone()
- Parent()

10



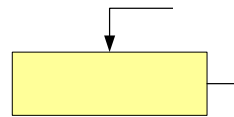
## Implementação do modelo

### ■ Ficheiros XML

- Utilização do DOM

### ■ BD Relacionais

- Uma **coluna** por cada campo
- Id, **ParentId**, HasChildren
- Relação **circular**
- Valores **nulos** nos campos opcionais
- Tabelas adicionais para **listas**



11



## Avaliação do modelo

### ■ Vantagens

- **Simple**s de implementar e manipular
- Permite a aplicação de **catamorfismos**
- Permite a **migração** de dados entre modelos físicos de armazenamento

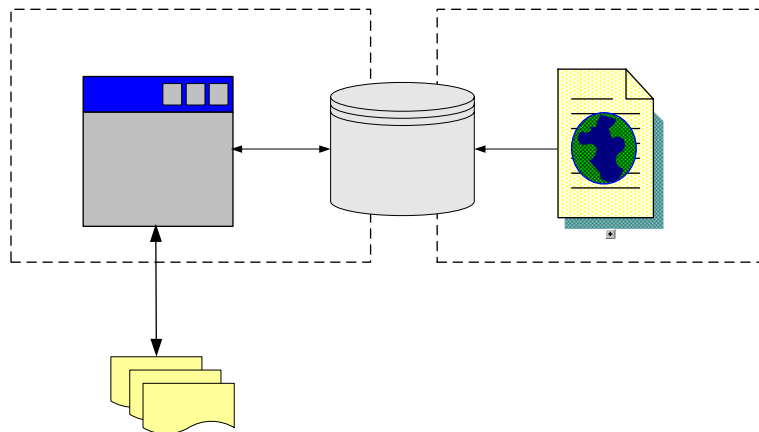
### ■ Desvantagens

- Não tira máximo partido das **capacidades** das BDR
- Encontrar um registo exige uma **travessia** na árvore inteira

12



## DigitArq [produto final]



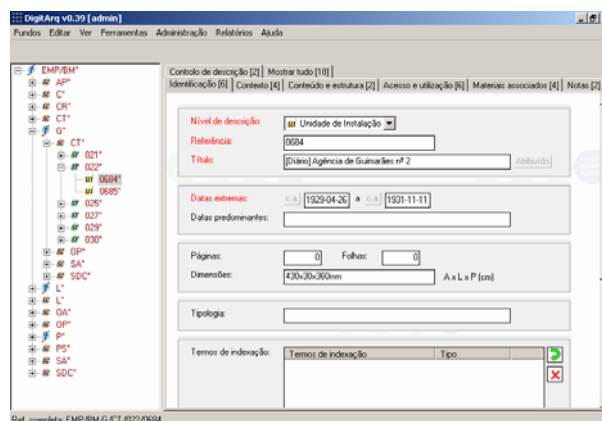
13



Software de Descrição Arquivística

## Software de descrição

- Apenas utiliza referências relativas
- Revisão automática da descrição
- Mecanismos de inferência



14



Base

## Conclusões

---

- Modelo **abstracto** para armazenamento de metainformação hierárquica
- **Simplicidade** e transparência
- Possibilidade de implementar sobre **outros modelos** físicos de armazenamento
- Facilidade em **migrar** dados entre modelos

15



## Trabalho futuro

---

- **Optimizar** o modelo para BDR
  - Stored procedures
- **Cache**
  - Minimizar o **tráfego** entre a aplicação e a BD
- **Pre-loading**
  - **Antever** a necessidade de informação

16





## Questões?

---

# Aquisição e Armazenamento de Metainformação no Contexto de um Arquivo

**Miguel Ferreira**  
mferreira@dsi.uminho.pt

*Co-Autoria*  
**José Carlos Ramalho**  
jcr@di.uminho.pt